

Processamento de Fala 2006/07

3º Mini-teste

14 de Dezembro de 2006

Identifique o seu enunciado colocando o seu nome e número de aluno no espaço reservado no final. Só são aceites respostas às questões de escolha múltipla assinaladas no local apropriado no final do enunciado quando este estiver identificado.

1. Diga se as seguintes afirmações são verdadeiras ou falsas:
 - (a) As palavras muito frequentes têm maior probabilidade de serem pronunciadas perto da forma canónica que as pouco frequentes.
 - (b) Usam-se modelos HMM de 3 estados para reconhecimento de palavras isoladas correspondentes aos meses do ano.
 - (c) No cálculo de medidas de confiança é frequente aproximar-se a probabilidade de uma observação não pertencer a uma dada classe pela soma das probabilidades de pertencer às N classes rivais mais perto.
 - (d) A utilização de modelos HMM em alinhamento forçado é útil como ponto de partida de sistemas de síntese de fala por concatenação de unidades de comprimento variável.
 - (e) A taxa de erro de um sistema de verificação do orador depende do número de oradores para os quais o sistema foi treinado.
 - (f) A perplexidade de um sistema de reconhecimento de fala contínua aumenta tipicamente ao passar de modelos de bigramas para modelos de trigramas.
2. Indique a gama típica do número de palavras de código usadas em reconhecimento de fala com distribuições discretas.
3. Como se designa o método de alisamento de modelos de língua que usa a probabilidade do difone se o número de ocorrências do trifone não atingir um dado limiar?
4. Que medidas se usam para aferir o desempenho de um sistema de detecção de palavras-chave?
5. Indique duas medidas de distância que explorem a parametrização eficiente da envolvente espectral feita por técnicas LPC.
6. Indique 3 tipos de características que utilizaria na construção de uma árvore de decisão para o alinhamento de corpora usando pronúncias alternativas.
7. Indique 3 causas que podem afectar negativamente o reconhecimento de um relato típico de futebol.
8. Considere o reconhecimento de uma locução de 2 segundos. Quantos são os parâmetros calculados na fase de parametrização acústica no total da locução? Indique os cálculos, especificando o significado de cada grandeza envolvida nesses cálculos.
9. Indique duas técnicas de adaptação a ambientes adversos utilizadas no 1º e 3º trabalhos de laboratório, respectivamente, e classifique cada uma como actuando no espaço do sinal, das características ou dos modelos.

10. Considere um HMM discreto de 2 estados, que modela a extracção de uma carta (com reposição) de 2 baralhos. As probabilidades de observação são as seguintes:

- Para o 1º baralho: todas as 52 cartas são equiprováveis
- Para o 2º baralho: Prob(ouros)=0.2; Prob(espadas)=0.3; Prob(copas)=0.2; Prob(paus)=0.3. Dentro de cada naipe, todas as 13 cartas são equiprováveis.

As probabilidades de transição entre estados são as seguintes:

- Prob(permanecer no 1º baralho)=0.6
- Prob(permanecer no 2º baralho)=0.7
- Prob(transitar do 1º baralho para o 2º baralho)=0.4
- Prob(transitar do 2º baralho para o 2º baralho)=0.3

A primeira extracção é feita do 1º baralho.

- (a) Qual a probabilidade de sair ouros na 2ª extracção?
- (b) Qual a probabilidade da 2ª extracção ser feita com o 2º baralho, sabendo que saiu o dez de paus.

(Nota: Indique apenas os cálculos a partir dos valores numéricos)

11. Considere um sistema de reconhecimento de fala contínua e vocabulário extenso, independente do orador. Para um dado segmento de fala, com algum ruído de fundo, a transcrição manual feita foi a seguinte:

A décima jornada da liga portuguesa de futebol começou em Aveiro. O Beira-Mar e o Vitória de Setúbal empataram a um golo no jogo que marcou a estreia de Carlos Carvalhal como treinador do Beira-Mar. Primeiro as boas-vindas ao novo treinador.

A transcrição automática produzida pelo reconhecedor foi:

A décima jornada da liga portuguesa de futebol começou em Aveiro o Beira-Mar e Vitória de. E empataram a um golo no jogo que marcou a estreia de Carlos. Carvalhal. Com o treinador do Beira-Mar. A primeira hoje boas-vindas ao novo treinador.

Ignorando a pontuação, complete os valores de H ("correct"), D ("deletions"), S ("substitutions"), I ("insertions"), N ("total"), %Corr, %Acc e %WER correspondentes.

Admita que cada palavra tem associada uma característica de pontuação *sem_pont / com_pont*. Calcule a exactidão (%ACC) do detector de pontuação automático. (*Transporte* para a transcrição automática a pontuação correcta e faça os cálculos sobre a transcrição automática.)

12. Considere o corpus composto pelas seguintes frases:

O Sebastião anda a pé. O Fernando anda de carro. O carro do Fernando não trabalha. O Sebastião não tem carro. O Fernando tem outro carro.

Considere a frase de teste:

O Sebastião não trabalha.

- (a) Calcule a probabilidade da frase usando um modelo de bigramas sem alisamento
- (b) Calcule a probabilidade da frase usando um modelo de bigramas com alisamento do tipo add-one
- (c) Calcule a probabilidade da frase usando um modelo de trigramas sem alisamento
- (d) Qual o número de ocorrências do(s) trigramas(s) mais frequentes neste pequeno corpus?

(Nota: Indique apenas os cálculos a partir dos valores numéricos)

Respostas

Nome:	
Número:	

1. (1,8 val.)

a	b	c	d	e	f

2 a 8 (1 / 1 / 1 / 1 / 1,2 / 1,2 / 1,5 valores)

2	
3	
4	
5	
6	
7	
8	

9. (1,8 val.)

técnica	espaço

10. (2,5 val.)

a)	
b)	

11. (3,0 val.)

H	D	S	I	N	%Corr	%Acc	%WER	%Acc-Pont

12. (3,0 val.)

a)	
b)	
c)	
d)	